Conformers of Gaseous Protonated Glycine

KUI ZHANG, ALICE CHUNG-PHILLIPS

Department of Chemistry, Miami University, Oxford, Ohio 45056

Received 16 March 1998; accepted 29 July 1998

ABSTRACT: *Ab initio* geometry optimizations were performed on gaseous protonated glycine using the second-order Møller–Plesset perturbation theory with the 6-31G*, 6-31G**, 6-31 + G**, and 6-311 + G** basis sets. Eight energy minima and 12 saddle points in the low-energy region of the electronic potential energy surface were characterized. The global minimum was an amino N-protonated conformer containing an ionic H bond between the $-NH_3^+$ and O=C< groups. The lowest energy O-protonated conformer was stabilized by a conjugative attraction between the nitrogen lone-pair electrons and the positively charged planar fragment $-C(OH)_2^+$. Relative electronic energies of the nine N- and 11 O-protonated species fall in the ranges of 0–10 and 30–40 kcal mol^{-1}. At room temperature the equilibrium distribution contained the most stable N-protonated conformer almost exclusively. Additional subjects for investigation include the effects of basis set and electron correlation on the predicted structures, nonbonded interactions that influence the relative stability of protonated conformers, conformational interconversions based on intrinsic reaction coordinate calculations, and kinetic pathways for protonation and associated changes in Gibbs free energy. The work provides geometric, energetic, and thermodynamic data pertinent to the study of gas-phase ion chemistry of amino acids and peptides. © 1998 John Wiley & Sons, Inc. J Comput Chem 19: 1862–1876, 1998

Keywords: protonation of glycine; *ab initio* calculations; conformational analysis; gas-phase basicity; amino acids

Correspondence to: A. Chung-Phillips; e-mail: ap32chmf@miamiu.muohio.edu

Contract/grant sponsors: Miami University Computing and Information Services, Ohio Supercomputer Center

Contract/grant sponsor: National Institute of General Medical Sciences; contract/grant number: R15-GM52670-01

This article includes Supplementary Material available from the authors upon request or via the Internet at ftp.wiley.com/public/journals/jcc/suppmat/19/1862 or http://journals.wiley.com/jcc/

CCC 0192-8651 / 98 / 161862-15

Introduction

Amino acids and peptides form the building blocks of proteins. Among biomolecules, peptides are unique in having both seemingly infinite conformational possibilities and a broad spectrum of functional groups. To gain an understanding of the biological activity of a peptide, it is often necessary to have 3-dimensional structural information on simpler species such as the isolated amino acids.[1,2]

In the gas phase, protonation is the dominant ionization pathway during the analysis of peptides by mass spectrometry,[3] which is one of the major experimental means to study biomolecules. The location of the proton affects the fragmentation patterns of the protonated peptide ions and consequently the structural information obtained from mass spectroscopy. Recently, the mass spectrometric measurements of the gas-phase basicities (GBs) of amino acids, simple peptides, and small carbohydrates by Cassady and coworkers,[4–10] complemented with *ab initio* studies from the present authors,[4,6–8,11] added a rich source of information to the study of protonations in these molecular systems.

The purpose of this research was to carry out a conformational analysis of the protonated glycine in the gas phase using the high-level *ab initio* approach. As the simplest amino acid, neutral glycine has been extensively studied by theoretical and experimental methods[12–26]; but relatively little attention has been given to protonated glycine.[4,7,27–30] The most comprehensive work on protonated glycine was done by Jensen[29] at the Hartree–Fock (HF) and second-order Møller–Plesset perturbation (MP2) levels of theory with the 6-31G* basis set: nine symmetric minimum-energy structures including three amino N- and six carbonyl O-protonated species were identified by geometry optimizations at the HF/6-31G* level, and relative energies were calculated at the MP2/6-31G* level. Recently, one asymmetric and one symmetric structure associated with the most stable N-protonated glycine were optimized by Yu et al.[30] at the MP2/6-31G* level and energies were determined at the G2(MP2) level. To our knowledge, there has not been a comprehensive study of protonated glycine involving geometry optimizations at the post-HF level using basis sets larger than 6-31G*. Such a study is necessary to better characterize protonated glycine, because it has been well documented for the neutral glycine that electron correlation has a great impact on the predicted structures[14,15,17,20,26] and that an expansion of the basis set may produce asymmetric geometries for certain energy minima.[11,12]

We present in this article the results of a study on protonated glycine conformers obtained from geometry optimizations using the MP2 theory over several double and triple split-valence sets including polarization and diffuse functions. The *ab initio* levels of choice are not as high as those employed by Császár[14,23] and Hu et al.[20] in their respective structural and conformational analysis of glycine, but the results are expected to be sufficiently accurate for the intent of this protonated study. We had four specific aims: to deduce the effects of basis set and electron correlation on the calculated protonated structures, to elucidate the relative stability of protonated conformers based on nonbonded interactions, to characterize the low-energy region of the electronic potential energy surface (PES) of protonated glycine, and to postulate simple kinetic pathways for protonation and to estimate the Gibbs free-energy changes associated with the proposed protonation reactions.

Computational Methods

Ab initio MO calculations were carried out using the Gaussian 92/94 program.[31,32] Structures of protonated glycine were fully optimized at the MP2/6-31G*, MP2/6-31G**, MP2/6-31 + G**, and MP2/6-311 + G** levels. Searches were made for the stationary points in the PES generated by internal rotation about the C—C, C—N, and C—O bonds; relevant dihedral angles are shown in Figure 1. Each energy minimum is located by the usual optimization procedure for an equilibrium structure. In searching for a particular saddle point the relevant dihedral angles are set near the idealized values for the intended transition structure before employing the optimization procedure for the transition state (TS). Note that each stationary point is verified by the calculated harmonic vibrational frequencies as an energy minimum (all positive frequencies) or a saddle point (with only one imaginary frequency). At the highest level, MP2/6-311 + G**, the structures were obtained by using analytic HF force constants and tight convergence cutoffs in geometry optimizations. Additional calculations were performed to obtain (a)

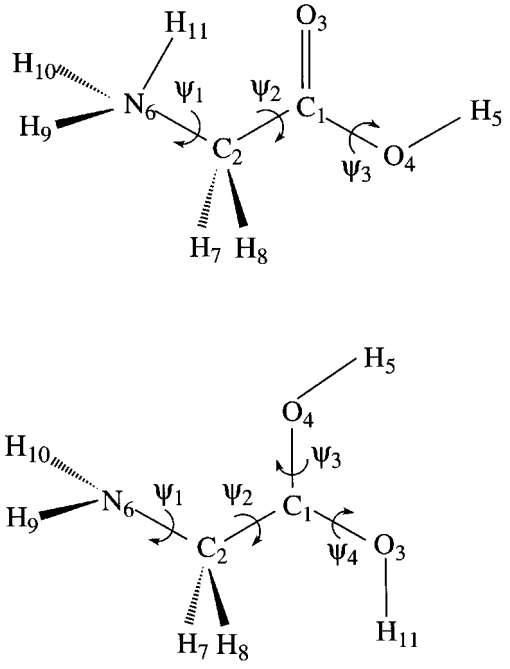

FIGURE 1. Definitions of dihedral angles ψ_1, ψ_2, ψ_3, and ψ_4. For the N-protonated glycine (top), $\psi_1 = H_{11}—N_6—C_2—C_1$, $\psi_2 = N_6—C_2—C_1—O_3$, and $\psi_3 = H_5—O_4—C_1—O_3$. For the O-protonated glycine (bottom), $\psi_1 = H_{10}—N_6—C_2—C_1$, $\psi_2 = N_6—C_2—C_1—O_3$, $\psi_3 = H_5—O_4—C_1—O_3$, and $\psi_4 = H_{11}—O_3—C_1—C_2$. The symmetric structures 1*n* (top) and 4*n* (bottom) are shown.

energies at the fourth-order MP treatment, MP4/6-31G*//MP2/6-31G*, for all structures; (b) harmonic vibrational frequencies at both the MP2/6-31+G** and MP2/6-311+G** optimized levels for all structures; and (c) intrinsic reaction coordinates (IRCs) at the MP2/6-31G* optimized level for all saddle points. Note the levels chosen for (a) and (c) are expected to yield reasonably accurate values at reasonable cost for the respective properties. The correlation treatments MP2 and MP4 include all electrons.

Results and Discussion

PHYSICAL PROPERTIES

In this study several physical properties are computed for the individual protonated glycine conformers to estimate their relative stability, populations, and contributions to the GB of glycine. The relevant equations are presented below.

The Gibbs free energy G at temperature T for a pure ideal gas is calculated as

$$G = E_e + E_{ZP} + (E - E_0) + RT - TS \quad (1)$$

where E_e is the electronic energy, E_{ZP} is the zero-point energy, $(E - E_0)$ is the thermodynamic internal energy, and S is the entropy. To characterize the PES, we first employ E_e to locate the minima and transition states and next determine their relative energies by including vibrational motion at 0 K: $E_0 = E_e + E_{ZP}$. When used for the calculations of G and E_0, the E_{ZP} values are scaled by appropriate factors.[33,34]

For ease of comparison, we express the property $M(i)$ for conformer i relative to $M(1)$ of the lowest energy conformer 1 in the form of the relative property $\Delta M(i)$: $\Delta M(i) = M(i) - M(1)$. At 298.15 K and 1 atm, the relative Gibbs free energy of conformer i is

$$\Delta G(i) = \Delta E_e(i) + \Delta E_{ZP}(i) + \Delta(E - E_0)(i) - (298.15\ \mathrm{K})\Delta S(i), \quad (2)$$

relative to $\Delta G(1) = 0.00$. The corresponding equilibrium population $p(i)$ is estimated by a Boltzmann expression[13]

$$p(i) = \frac{e^{-\Delta G(i)/RT}}{\sum_i e^{-\Delta G(i)/RT}}. \quad (3)$$

For the protonation reaction of glycine (Gly) at 298.15 K and 1 atm, $Gly + H^+ \rightarrow GlyH^+$, the two principal thermodynamic parameters are GB and proton affinity (PA) defined as the negative of the Gibbs free-energy change and enthalpy change: $GB = -\Delta G_r$ and $PA = -\Delta H_r$. The two parameters are related by the entropy term: $GB = PA + T\Delta S_r$. For the protonation reaction *rk* that involves the *n*th neutral conformer as the reactant and the *p*th protonated conformer as the product,

$$\mathrm{Gly}(n) + \mathrm{H}^+ \rightarrow \mathrm{GlyH}^+(p), \quad (\mathrm{rk})$$

the thermodynamic parameters may be related to those of reaction *r*1 that specifically involves the respective lowest energy conformers,

$$\mathrm{Gly}(1) + \mathrm{H}^+ \rightarrow \mathrm{GlyH}^+(1). \quad (\mathrm{r1})$$

Thus, the GB for *rk*, GB(*rk*), may be related to the GB for *r*1, *GB*(*r*1), via the relation

$$\Delta\mathrm{GB}(rk) = \mathrm{GB}(rk) - \mathrm{GB}(r1). \quad (4)$$

Utilizing the available $\Delta G(i)$ of eq. (2), the relative GB for rk, $\Delta GB(rk)$, is directly deduced as

$$\Delta GB(rk) = -[\Delta G(p) - \Delta G(n)], \quad (5)$$

where $\Delta G(n) = G(n) - G(1)$ is for the Gly conformers and $\Delta G(p) = G(p) - G(1)$ is for the $GlyH^+$ conformers. The $GB(r1)$, which corresponds to the most probable protonation reaction due to the abundance of Gly(1) and $GlyH^+(1)$ at ordinary temperature, is calculated as[11]

$$GB(r1) = -\left[\Delta E_e(r1) + \Delta G_{therm}(r1) + 6.28 \text{ kcal mol}^{-1}\right] \quad (6)$$

where

$$\Delta G_{therm}(r1) = \Delta E_{ZP}(r1) + \Delta(E - E_0)(r1) - (298.15 \text{ K})\Delta S(r1) \quad (7)$$

In this context, the relative property is $\Delta M(r1) = M[GlyH^+(1)] - M[Gly(1)]$. For the calculation of $PA(r1)$ from $GB(r1)$,

$$T\Delta S(r1) = (298.15 \text{ K})\{S[GlyH^+(1)] - S[Gly(1)]\} - 7.76 \text{ kcal mol}^{-1} \quad (8)$$

MP2 OPTIMIZATIONS

Császár showed previously that molecular properties of the glycine conformers calculated at the MP2/6-311++G** optimized level are in good agreement with experiments.[14] We found in a recent study on the GB of glycine that the calculated geometries and relative energies of Gly(1) and $GlyH^+(1)$ are virtually the same at the MP2/6-311+G** and MP2/6-311++G** optimized levels.[11] In other words, the addition of s diffuse functions on hydrogen atoms to the triple split-valence basis 6-311+G** has an insignificant effect on the calculated structures. Consequently we chose 6-311+G** as the upper basis limit to study the effect of different basis sets on the MP2 structures. Although the more balanced and complete basis sets optimized for correlated calculations by Dunning and colleagues[35] may provide more definitive electronic structures, it is not practical at present to use a basis set comparable to aug-cc-pVTZ on the $GlyH^+$ conformers.[11]

Geometry optimizations at the MP2/6-31G* level produced an initial set of 18 electronic structures for the protonated glycine, all of which were reproduced in subsequent optimizations at the MP2/6-31G**, MP2/6-31+G**, and MP2/6-311+G** levels. However, the two larger basis sets introduced new asymmetric structures corresponding to the second and third most stable energy minima as a result of including sp diffuse functions on the nitrogen and oxygen atoms. The 20 structures optimized at the highest level, MP2/6-311+G**, are shown in order of increasing energy for the nine N-protonated ($1m, \ldots, 3t$) and 11 O-protonated ($4m, \ldots, 8t$) glycine in Figures 2 and 3; the point-group symmetry of each structure is indicated within parentheses. The corresponding geometrical parameters and energies are provided in Tables I–III. These structures correspond to eight energy minima ($1m$–$8m$) and 12 transition states; the latter are further characterized as satellite structures of the first four minima ($1n$–$4n$) and transition structures for internal rotation and internal proton transfer ($1t$–$8t$). Here, each TS is a saddle point on the PES.

Unless stated otherwise, the MP2/6-311+G** results are used to represent the protonated structures. Protonation of glycine at the amino nitrogen or carbonyl oxygen site leads to the formation of a new ionic N—H or O—H bond; consequently, the two groups of structural isomers produced are

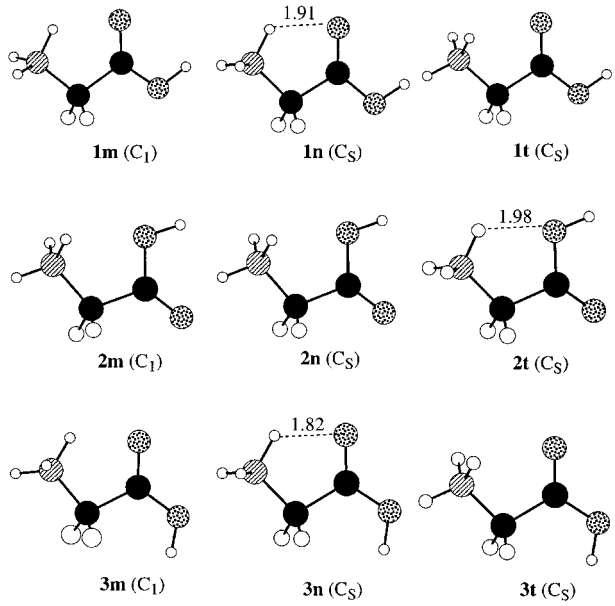

FIGURE 2. MP2/6-311+G** structures for the N-protonated glycine. The shortest H···O distances in angstroms for groups 1, 2, and 3 are shown by dashed lines.

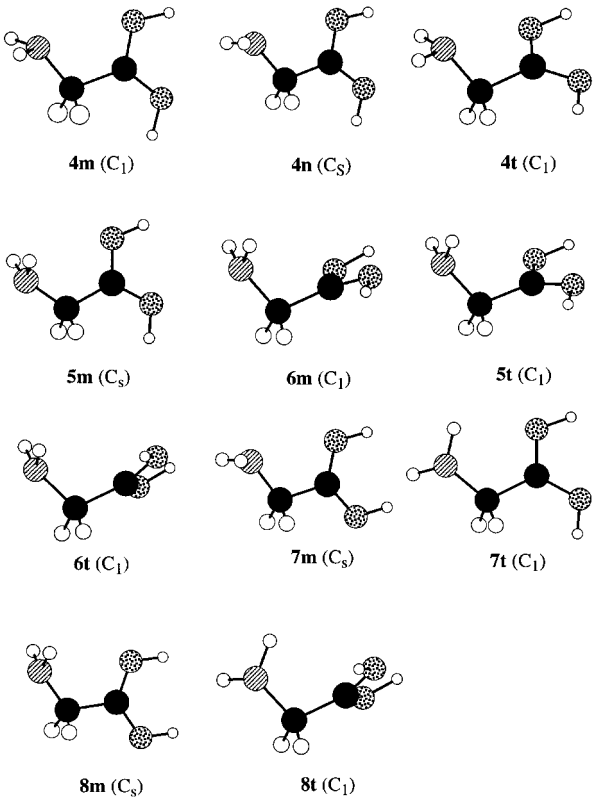

FIGURE 3. MP2 / 6-311 + G** structures for the O-protonated glycine.

the N- and O-protonated glycine. The O-protonated species are shown to be 20–30 kcal mol^{-1} higher in energy than the N-protonated species (Table III), which is consistent with the magnitude of decrease in the PAs observed for carboxylic acids relative to those of the amines of comparable mass and configuration. For example, the experimental PAs of 2-methyl-1-propanamine ($C_4H_{11}N$) and propanoic acid ($C_3H_6O_2$) are 220.2 and 190.6 kcal mol^{-1}, respectively.[36] Changes in geometry and energy induced by protonation may be examined by comparing the calculated values for the protonated glycine in Tables I–III with Császár's values for the neutral glycine in his tables I and III of ref. 14.

To gain a better understanding of the properties of protonated ($GlyH^+$) conformers relative to those of the neutral (Gly) conformers the physical properties including dipole moments (μ), relative electronic energies (ΔE_e), relative Gibbs free energies at 298.15 K (ΔG), and equilibrium populations (p) at 298.15 K and 1 atm for the eight lowest energy conformational minima in the PESs of Gly and $GlyH^+$ are provided in Table IV for reference. The μ and ΔE_e data for the neutral conformers Ip, ..., VIIIn were computed by Császár[14] at the MP2/6-311++G** level; those for the protonated conformers 1m, ..., 8m were calculated at the MP2/6-311+G** level (cf. Table III). For the μ of Gly we included a corrected value for Ip and a new value for IIIn (Table IV). The μ of $GlyH^+$ is reported for general information because it is related to the overall charge distribution and is often used to characterize a calculated structure. Thermal contributions to ΔG [i.e., the ΔE_e, ΔE_{ZP}, $\Delta(E - E_0)$, and ΔS terms of eq. (2)] were determined at the MP2/6-31G* level and are reported in Table S-I as Supplementary Material. Note the unscaled ΔE_{ZP} are listed in Table S-I, whereas the scaled ΔE_{ZP} are used for ΔG calculations in Table IV. The scale factor of 0.9646 was derived by Pople et al.[33] for the MP2(fu)/6-31G* level, which is the same factor used in our recent *ab initio* calculations of the gas-phase basicity of glycine.[11] Structures of six of the eight neutral conformers are reproduced from ref. 11 in the top two rows of Figure 4. (Structures of all calculated neutral glycine species are shown in ref. 14.) Recently, quantitative measurements were made by means of infrared,[22] jet,[24] and electron momentum[25] spectroscopy to identify all the neutral glycine conformers proposed by theory. It is hoped that similar experimental measurements will be made on the protonated glycine conformers proposed by this and other theoretical studies. In this regard the theoretical data on μ and ΔE_e in Table IV, which show a wide separation in values between the N- and O-protonated species, may prove useful to the design and analysis of experiments.

The most stable N-protonated conformer 1m (cf. 1n at the top of Fig. 1) has an asymmetric geometry that possesses an ionic H bond between the ammonium H_{11} and carbonyl O_3 atoms, $NH^+ \cdots OC$, and assumes a cis form for the carboxyl group $O_3—C_1—O_4—H_5$. The conformation suggests the existence of a five-membered ring $H_{11}—N_6—C_2—C_1—O_3$ closed by $H_{11} \cdots O_3$. This protonated species may be postulated to have formed from an N protonation of the most abundant neutral conformer Ip (65%, Table IV) that leads to the transition state 1t before falling into the minimum 1m. Primary geometric changes take place at the N terminus due to the addition of a proton to form the ammonium group (e.g., lengthening of the C—N and N—H bonds) and at the atoms affected by the ring formation

TABLE I.
Geometrical Parameters for N-Protonated Glycine Structures Optimized at MP2 / 6-311 + G Level.**

Parameter	1*m*	1*n*	1*t*	2*m*	2*n*	2*t*	3*m*	3*n*	3*t*
$C_2—C_1$	1.525	1.527	1.522	1.524	1.524	1.526	1.541	1.541	1.535
$O_3—C_1$	1.210	1.211	1.209	1.195	1.195	1.195	1.208	1.208	1.205
$O_4—C_1$	1.319	1.317	1.322	1.352	1.351	1.352	1.320	1.319	1.325
$H_5—O_4$	0.972	0.972	0.972	0.972	0.972	0.971	0.967	0.967	0.966
$N_6—C_2$	1.498	1.500	1.498	1.504	1.504	1.512	1.497	1.497	1.497
$H_7—C_2$	1.090	1.090	1.091	1.090	1.091	1.090	1.092	1.091	1.092
$H_8—C_2$	1.091	1.090	1.091	1.092	1.091	1.090	1.091	1.091	1.092
$H_9—N_6$	1.023	1.022	1.024	1.024	1.024	1.023	1.023	1.022	1.024
$H_{10}—N_6$	1.024	1.022	1.025	1.025	1.025	1.023	1.022	1.022	1.026
$H_{11}—N_6$	1.033	1.041	1.025	1.026	1.025	1.030	1.045	1.048	1.026
$H_{11}\cdots$ (O_3 or O_4)[a]	2.074	1.910	2.503	2.284	2.461	1.977	1.854	1.822	2.447
$O_3—C_1—C_2$	121.0	120.9	121.2	122.8	122.5	122.2	118.8	118.9	119.1
$O_4—C_1—O_2$	110.9	111.0	110.5	109.8	110.2	110.7	117.8	117.8	117.4
$H_5—O_4—C_1$	109.2	109.4	109.1	108.9	108.8	109.3	112.8	112.8	113.0
$N_6—C_2—C_1$	105.9	105.6	106.5	111.5	111.8	112.2	104.9	104.8	106.0
$H_7—C_2—C_1$	112.5	110.9	111.5	110.0	109.6	108.9	110.4	111.5	112.1
$H_8—C_2—C_1$	109.9	110.9	111.5	109.2	109.5	108.9	112.7	111.5	112.1
$H_9—N_6—C_2$	112.9	112.8	112.0	111.6	111.5	112.0	112.6	113.1	112.3
$H_{10}—N_6—C_2$	111.7	112.8	110.5	111.4	111.0	112.0	113.4	113.1	110.3
$H_{11}—N_6—C_2$	107.4	105.5	110.5	110.3	111.0	109.1	104.4	103.9	110.3
$O_4—C_1—C_2—O_3$	178.9	180.0	180.0	−178.5	180.0	180.0	−178.9	180.0	180.0
$H_5—O_4—C_1—O_3$ (ψ_3)	1.4	0.0	0.0	−0.1	0.0	0.0	178.8	180.0	180.0
$N_6—C_2—C_1—O_3$ (ψ_2)	10.9	0.0	0.0	−166.2	180.0	180.0	−6.6	0.0	0.0
$H_7—C_2—C_1—N_6$	120.3	119.2	118.6	120.9	120.2	120.8	117.4	118.6	118.0
$H_8—C_2—C_1—N_6$	−117.4	−119.2	−118.6	−119.2	−120.2	−120.8	−119.6	−118.6	−118.0
$H_9—N_6—C_2—C_1$	−151.3	−118.7	180.0	−174.0	180.0	−119.5	−104.2	−118.3	180.0
$H_{10}—N_6—C_2—C_1$	86.7	118.7	58.3	65.2	58.9	119.5	132.5	118.3	57.9
$H_{11}—N_6—C_2—C_1$ (ψ_1)	−29.6	0.0	−58.3	−52.8	−58.9	0.0	12.4	0.0	−57.9

See Figure 1 (top) for atom labels and Figure 2 for structures. Bond length A—B is in angstroms; bond angle A—B—C and dihedral angle A—B—C—D are in degrees, where A—B—C—D is measured clockwise from A—B to C—D viewed in the B—C direction.

[a]Nonbonded distance (Å) between H_{11} and O_3 for 1*m*, 1*n*, 1*t*, 3*m*, 3*n*, and 3*t* or between H_{11} and O_4 for 2*m*, 2*n*, and 2*t*.

due to hydrogen bonding (e.g., contraction of the C—C=O, C—C—N, and C—N—H bond angles).

The most stable O-protonated conformer 4*m* (cf. 4*n* at the bottom of Fig. 1) is also asymmetric; it contains an ionic planar fragment $H_5—O_4—C_1—O_3—H_{11}$ in a cis-trans form, and the amino hydrogen atoms H_9 and H_{10} point away from the $C_1—O_4$ bond. A simple kinetic path for the formation of 4*m* from the lower energy neutral conformers is not forthcoming; however, a one-step O protonation on the scarcely populated V*n* (1%, Table IV) may be envisioned. Protonation induces an equilibration of the two C—O bonds and C—C—O angles expected for the π-electron conjugation in the planar fragment. There is also a decrease in the C—N bond length that indicates partial delocalization of the nitrogen lone pair into the nearly planar heavy-atom framework formed by N_6, C_2, C_1, O_4, and O_3.

BASIS SETS AND ELECTRON CORRELATION

Tables S-II and S-III document the changes in geometry for the 20 conformers 1*m* −8*t* as a result of basis expansion from MP2/6-31G*, MP2/6-31G**, and MP2/6-31 + G** to MP2/6-311 + G**. The MP2 geometrical parameters for the global minimum 1*m* in Table S-II show very small changes from one basis set to another: bond lengths within 0.01 Å and bond and dihedral angles within 1° except for those associated with $—NH_3^+$ (cf. values of ψ_1). For the remaining 19 conformers, all parameter values vary within a very small range

TABLE II.
Geometrical Parameters for O-Protonated Glycine Structures Optimized at MP2 / 6-311 + G Level.**

Parameter	4*m*	4*n*	4*t*	5*m*	6*m*	5*t*	6*t*	7*m*	7*t*	8*m*	8*t*
$C_2—C_1$	1.508	1.510	1.504	1.513	1.506	1.510	1.509	1.506	1.488	1.510	1.490
$O_3—C_1$	1.289	1.289	1.284	1.287	1.287	1.286	1.286	1.275	1.286	1.280	1.286
$O_4—C_1$	1.268	1.268	1.272	1.273	1.277	1.277	1.276	1.287	1.275	1.284	1.274
$H_5—O_4$	0.978	0.978	0.977	0.978	0.977	0.977	0.978	0.973	0.978	0.974	0.978
$N_6—C_2$	1.429	1.428	1.447	1.420	1.442	1.439	1.438	1.430	1.437	1.420	1.467
$H_7—C_2$	1.101	1.100	1.102	1.099	1.092	1.094	1.091	1.098	1.103	1.098	1.091
$H_8—C_2$	1.099	1.100	1.093	1.099	1.093	1.094	1.095	1.098	1.101	1.098	1.097
$H_9—N_6$	1.010	1.010	1.012	1.011	1.012	1.012	1.011	1.010	1.011	1.010	1.013
$H_{10}—N_6$	1.011	1.010	1.013	1.011	1.012	1.012	1.010	1.010	1.011	1.010	1.013
$H_{11}—O_3$	0.972	0.972	0.972	0.973	0.972	0.972	0.972	0.974	0.973	0.974	0.974
$O_3—C_1—C_2$	122.8	122.8	125.3	123.7	124.8	125.1	124.6	119.0	123.6	118.3	122.8
$O_4—C_1—O_2$	119.7	119.7	117.1	119.0	117.7	117.6	118.2	115.8	119.2	116.6	119.6
$H_5—O_4—C_1$	111.8	111.8	112.7	112.4	112.5	112.6	112.6	116.2	112.5	116.5	112.5
$N_6—C_2—C_1$	112.1	112.2	108.0	118.0	110.3	112.2	111.8	112.2	115.1	118.1	108.7
$H_7—C_2—C_1$	105.8	105.1	105.8	105.3	107.9	107.2	107.2	104.8	103.7	105.0	107.8
$H_8—C_2—C_1$	104.6	105.1	108.4	105.3	109.3	108.8	108.4	104.8	105.8	105.0	108.0
$H_9—N_6—C_2$	112.8	113.5	111.2	114.8	114.7	114.5	115.8	113.1	112.3	115.1	111.2
$H_{10}—N_6—C_2$	113.6	113.5	112.6	114.8	115.7	115.1	117.3	113.1	115.0	115.1	115.8
$H_{11}—O_3—C_1$	112.8	112.9	112.7	112.8	112.5	112.4	112.7	115.4	112.9	116.0	112.5
$O_4—C_1—C_2—O_3$	−179.3	180.0	179.1	180.0	−174.1	−175.7	−175.5	180.0	179.3	180.0	−176.9
$H_5—O_4—C_1—O_3$ (ψ_3)	0.0	0.0	−2.3	0.0	−0.9	−0.5	0.1	0.0	−0.3	0.0	0.1
$N_6—C_2—C_1—O_3$ (ψ_2)	−175.0	180.0	131.2	180.0	95.4	109.1	68.6	0.0	163.2	0.0	70.0
$H_7—C_2—C_1—N_6$	127.0	124.2	125.1	124.1	120.4	120.9	121.5	124.7	124.6	124.7	122.5
$H_8—C_2—C_1—N_6$	−121.2	−124.2	−119.9	−124.1	−121.5	−122.3	−121.7	−124.7	−124.1	−124.7	−120.5
$H_9—N_6—C_2—C_1$	−132.0	−116.7	−146.8	65.4	67.0	68.0	60.8	−117.3	131.3	66.0	123.1
$H_{10}—N_6—C_2—C_1$ (ψ_1)	102.2	116.7	90.8	−65.4	−65.6	−63.2	−76.4	117.3	5.3	−66.0	−1.0
$H_{11}—O_3—C_1—C_2$ (ψ_4)	1.6	0.0	−2.6	0.0	8.3	5.5	4.9	180.0	−2.9	180.0	4.6

See Figure 1 (bottom) for atom labels and Figure 3 for structures. See Table I for parameters.

TABLE III.
Relative Energies for Protonated Glycine Structures Optimized at MP2 / 6-31G*, MP2 / 6-31G, MP2 / 6-31 + G**, and MP2 / 6-311 + G** Levels.**

	MP2 / 6-31G*		MP2 / 6-31G**	MP2 / 6-31 + G**		MP2 / 6-311 + G**	
	ΔE_e	ΔE_e (MP4)	ΔE_e	ΔE_e	ΔE_{ZP}	ΔE_e	ΔE_{ZP}
1*m*	0.00	0.00	0.00	0.00	0.00	0.00	0.00
1*n*	0.02	0.02	0.01	0.11	−0.19	0.13	−0.24
1*t*	0.37	0.37	0.48	0.27	−0.06	0.17	0.01
2*m*	—	—	—	3.69	−0.08	3.53	0.05
2*n*	3.90	3.81	4.15	3.69	−0.11	3.58	−0.09
2*t*	5.38	5.28	5.60	5.19	−0.34	5.05	−0.38
3*m*	—	—	—	9.42	−0.45	9.11	−0.52
3*n*	9.92	9.69	9.50	9.43	−0.55	9.12	−0.65
3*t*	10.51	10.25	10.30	9.93	−0.36	9.56	−0.33
4*m*	33.67	33.00	30.50	30.17	−1.71	28.44	−1.55
4*n*	33.68	33.02	30.51	30.18	−1.86	28.46	−1.73
4*t*	34.29	33.56	31.19	30.52	−1.34	28.88	−1.22
5*m*	34.15	33.52	30.95	30.74	−1.56	29.05	−1.43
6*m*	35.74	35.06	32.59	32.03	−1.39	30.39	−1.16
5*t*	35.86	35.13	32.69	32.10	−1.39	30.42	−1.18
6*t*	36.29	35.67	33.12	32.36	−1.76	30.73	−1.35
7*m*	37.64	36.95	34.39	34.06	−1.96	32.66	−1.88
7*t*	39.07	38.43	35.80	34.93	−1.85	32.87	−1.73
8*m*	39.04	38.37	35.70	35.37	−1.84	34.00	−1.75
8*t*	40.58	39.88	37.47	36.61	−1.49	34.80	−1.41

See Figures 2 and 3. Electronic energy E_e, single-point E_e at MP4 / 6-31G* // MP2 / 6-31G* designated as E_e (MP4), and zero-point energy E_{ZP} are listed in Table S-IV. Energies listed in this table are relative to those of 1*m* (kcal mol^{-1}). The ΔE_{ZP} values have been scaled by 0.9608 for MP2 / 6-31 + G** and 0.9748 for MP2 / 6-311 + G** (ref. 34). At the MP2 / 6-31G* and MP2 / 6-31G** levels, 2*m* and 3*m* of C_1 symmetry are nonexistent while 2*n* and 3*n* of C_s symmetry become the energy minima.

among all four basis sets; this is shown in part by the dihedral angles (ψ_1, ψ_2, ψ_3, and ψ_4) for structures 1*n*−8*t* in Table S-III.

To show the influence of the different basis sets on the calculated energy, relative electronic energies (ΔE_e) for all 20 structures obtained from the four different basis sets are listed in Table III. To investigate the effect of electron correlation on energy, single-point MP4/6-31G* calculations were performed on the MP2/6-31G* geometries; results are shown in Table III as MP4. To gain a better estimate of energy at 0 K, relative zero-point energies (ΔE_{ZP}) at the MP2/6-31 + G** and MP2/6-311 + G** levels were also included in Table III, which were respectively scaled by the factors 0.9608 and 0.9748 derived by Scott and Radom[34] for the MP2(fc)/6-31G** and MP2(fc)/6-311G** levels. Directly calculated E_e and unscaled E_{ZP} employed for deriving the relative values in Table III are provided in Table S-IV.

The relative energies (ΔE_e) produced by the four MP2 optimized levels in Table III show essentially the same order of relative stability for the 20 structures. For the ΔE_e of each O-protonated structure, the 6-31G* value is over 3 kcal mol^{-1} higher than the 6-31G** value. From 6-31G** to 6-31 + G** the corresponding changes in ΔE_e are less than 1 kcal mol^{-1}. These results suggest that 6-31G* is a less satisfactory basis for energy than geometry and the addition of *p* polarization functions to the hydrogen atoms (e.g., 6-31G**) improves the relative energies. On going from 6-31 + G** to 6-311 + G**, the ΔE_e for each O-protonated species is reduced up to 2 kcal mol^{-1}; this implies that expansion from a double to a triple split-valence basis has a considerable effect on the calculated energies. These comparisons show that to obtain reasonable energies for the O-protonated species an extended basis such as 6-311 + G** is needed to represent ionicity associated with the highly electronegative oxygen atom.

The order of stability calculated at MP4 shows no change from MP2 based on the ΔE_e values in the first two columns of Table III. But there is a small decrease in the relative energy from MP2 to MP4 at 0.0−0.3 and 0.6−0.7 kcal mol^{-1} for the N-

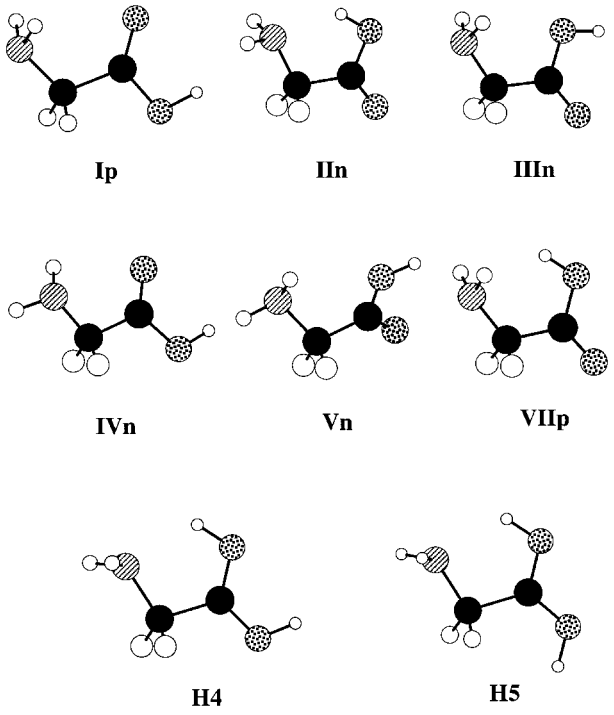

FIGURE 4. Selected structures of neutral (I*p*, II*n*, III*n*, IV*n*, V*n*, and VII*p*) and protonated (H4 and H5) glycine.

and O-protonated species, respectively. From these results we conclude that the MP expansion from MP2 to MP4 is rapidly converging and the second-order treatment (MP2) may indeed be satisfactory for the calculation of relative energies.

COMPARISON WITH HF STRUCTURES

The effect of including electron correlation in geometry optimizations can be deduced in part by comparing Jensen's HF/6-31G* to the present MP2/6-31G* structures in Table V. The HF structures H1, H2, H3, H6, H7, H8, and H9 (fig. 1 of ref. 29) correspond to the respective MP2 structures $1t$, $2n$, $3t$, $4n$, $5m$, $7m$, and $8m$; however, the HF structures H4 and H5 failed to surface from MP2 optimizations. From separate HF/6-31G* calculations starting with the MP2 structures missing from Jensen's list, we verified that $1n$ is an HF saddle point and $1m$, $2m$, $2t$, $3m$, $3n$, $4m$, $4t$, $6m$, $5t$, $6t$, $7t$, and $8t$ are nonexistent at the HF level.

There is a distinct symmetry difference between the HF and MP2 structures for four of the energy minima. The HF structures H1, H2, H3, and H6 are symmetric whereas the MP2 structures for the

TABLE IV.
Dipole Moments, Relative Electronic Energies, Relative Gibbs Free Energies, and Equilibrium Populations for Low-Energy Conformers of Glycine and Protonated Glycine.

	Symmetry	μ	ΔE_e	ΔG	p
		Glycine (Gly)			
I*p*	C_s	1.30	0.00	0.00	64.7
II*n*	C_1	5.59	0.51	1.26	7.7
III*n*	C_1	1.90	1.45	0.70	19.8
IV*n*	C_1	2.06	1.26	1.36	6.5
V*n*	C_1	2.41	2.21	2.34	1.2
VI*p*	C_s	2.95	5.71	5.57	—
VII*p*	C_s	4.10	6.95	6.88	—
VIII*n*	C_1	4.31	7.09	7.16	—
		Protonated Glycine ($GlyH^+$)			
1*m*	C_1	6.33	0.00	0.00	99.4
2*m*	C_1	8.17	3.53	3.03	0.6
3*m*	C_1	8.79	9.11	7.82	—
4*m*	C_1	3.14	28.44	26.31	—
5*m*	C_s	3.04	29.05	27.58	—
6*m*	C_1	2.70	30.39	28.94	—
7*m*	C_s	3.49	32.66	29.75	—
8*m*	C_s	5.36	34.00	32.23	—

See Figures 2–4 and figure 4 of ref. 11. Dipole moments are μ in debyes; energies relative to those of the lowest energy conformer I*p* or 1*m* are in kilocalories per mole. The μ and ΔE_e values for Gly are those of Császár[14] at the MP2/6-311++G** level; the μ values of I*p* and III*n* are corrected by the present authors. The μ and ΔE_e values for $GlyH^+$ are calculated at the MP2/6-311+G** level (cf. Table III). ΔG (kcal mol^{-1}) is the Gibbs free energy at 298.15 K and 1 atm, $\Delta G = \Delta E_e + \Delta G_{therm}$, where ΔG_{therm} is calculated at the MP2/6-31G* level and is listed in Table S-I. The ΔE_{ZP} values used for ΔG_{therm} have been scaled by 0.9646 (ref. 33). p(%) is the equilibrium population at 298.15 K and 1 atm; see eq. (3).

TABLE V.
Relative Electronic Energies for Selected Protonated Glycine Structures Optimized at HF / 6-31G* and MP2 / 6-31G* Levels.

HF	HF / 6-31G* // HF / 6-31G*	MP2 / 6-31G* // HF / 6-31G*	MP2	MP2 / 6-31G* // MP2 / 6-31G*
H1	0.0	0.0	1t	0.0
H2	4.6	3.7	2n	3.5
H3	10.8	10.1	3t	10.1
H4	13.8	17.6	—	—
H5	19.1	22.5	—	—
H6	26.5	32.7	4n	33.3
H7	27.1	33.2	5m	33.8
H8	30.2	36.7	7m	37.3
H9	31.6	38.1	8m	38.7

The values listed are ΔE_e (kcal mol^{-1}). Those calculated at the HF / 6-31G* geometries are those of Jensen.[29] The HF structures H4 and H5 (Fig. 4) have no MP2 counterparts.

same minima $1m$–$4m$ are asymmetric. Considering the closeness in energy between $1m$–$4m$ and their respective symmetric neighbors $1t$, $2n$, $3t$, and $4n$ that resemble the HF minima (cf. Table III), the symmetry difference is physically insignificant.

The significant difference lies in the presence of the HF structures H4 and H5 shown in the bottom row of Figure 4. The MP2 optimization using either the H4 or H5 geometry as the initial geometry always converged to the N-protonated species $1n$ or $3n$ because of proton migration (*vide infra*). These results suggest that electron correlation enhances the ionic H bond between the amino N and hydroxyl O, $H_2N\cdots H^+O$, to the extent that it self-destructs in the protonated glycine. [Recall that the neutral H bond, $H_2N\cdots HO$ in IIn of Fig. 4 was suggested as the strongest H bond that exists in glycine and alanine.[20,37]] Or rather, the HF level of theory grossly underestimates the strength of ionic hydrogen bonding.

Table V contains the comparisons of the MP2 and HF relative electronic energies (ΔE_e) for several structures at the HF/6-31G*//HF/6-31G*, MP2/6-31G*//HF/6-31G*, and MP2/6-31G*//MP2/6-31G* levels. Again, the order of stability is consistent among the three levels. However, on going from HF/6-31G* to MP2/6-31G*, both evaluated at the HF/6-31G* optimized geometries, ΔE_e of the O-protonated species is increased by about 6 kcal mol^{-1} due to electron correlation. Further improvement to the MP2/6-31G* optimized geometries leads to an additional small increase of 0.6 kcal mol^{-1} for each O-protonated species. These results suggest that MP2/6-31G*//HF/6-31G* calculations are indeed able to provide acceptable relative energies at considerable savings.[29]

N-PROTONATED STRUCTURES

The N-protonated conformers are separated into three groups by relative energies (last column of Table III): $1m$, $1n$, and $1t$ ($\sim$0 kcal mol^{-1}); $2m$, $2n$, and $2t$ (3–5 kcal mol^{-1}); and $3m$, $3n$, and $3t$ (9–10 kcal mol^{-1}). Groups 1 and 2 are lower in energy than group 3 because the cis form of —COOH is more stable than the trans form as a result of a strong attraction between the hydroxyl H and carbonyl O, $OH\cdots OC$. Group 1 is lower than group 2 because $NH^+\cdots OC$ is a stronger ionic H bond than $NH^+\cdots OH$, one between an ammonium H and the hydroxyl O. Again, this is consistent with the observed greater PA for 2-butanone (C_4H_8O) than 2-butanol ($C_4H_{10}O$) at 197.8 versus 194.8 kcal mol^{-1}, respectively.[36] From the relative energies of the three groups, we conclude that the hydrogen attractions follow the order $OH\cdots OC > NH^+\cdots OC > NH^+\cdots OH$. Strictly speaking, $OH\cdots OC$ is not an H bond because the $O—H\cdots O$ angle is not larger than 90°, a criterion widely accepted for an H bond.[38] For an H bond the shorter the $H\cdots O$ distance, the stronger the hydrogen bonding. The shortest $H\cdots O$ distance in each N-protonated conformer, $H_{11}\cdots O_3$ for $NH^+\cdots OC$ or $H_{11}\cdots O_4$ for $NH^+\cdots OH$, are provided in Table I for reference. Note the shortest $H\cdots O$ distances for groups 1, 2, and 3 are 1.91, 1.98, and 1.82 Å for $1n$, $2t$, and $3n$, respectively, which supports the deduced H-bond strength

$NH^+ \cdots OC > NH^+ \cdots OH$, regardless of whether —COOH is in a cis or trans form (see Fig. 2). As the H···O distance increases, it no longer serves as a sensitive indicator for H-bond strength (e.g., the bifurcated H-bond distances vary in the narrow range of 2.4–2.5 Å for all three groups).

The relative stability of the nine N-protonated structures $1m$–$3t$ may be explained by the relative strength of hydrogen bonding attraction (HBA) versus vicinal steric repulsion (VSR). Specifically, hydrogen bonding is between an N-terminus hydrogen and a carboxyl oxygen atom, O···H, in the 1,5 positions and steric repulsion includes mainly those among atoms bonded to the nitrogen and α carbon such as H···H and H···C repulsions in the 1,4 positions. The geometrical positions of these atoms are such that an increase in HBA is accompanied by an increase in VSR. As a result of interplay between these two opposing interactions, three types of geometry appear (Fig. 2). The first is the symmetric transition state $1t$, $2n$, or $3t$ that has a bifurcated H bond and a staggered ammonium conformation; this represents the lowest HBA and VSR. The second is the symmetric transition state $1n$, $2t$, or $3n$ that has the shortest H bond and an eclipsed ammonium conformation, resulting in the strongest HBA and VSR. The last is the asymmetric energy minimum $1m$, $2m$, or $3m$ that achieves the optimal balance between HBA and VSR. Note that $1m$, $2m$, and $3m$ are slightly lower in energy than their respective satellite structures $1n$, $2n$, and $3n$ (< 0.1 kcal mol^{-1}) after optimally balancing hydrogen bonding attraction with steric repulsion by small changes in ψ_1 and ψ_2.

O-PROTONATED STRUCTURES

It is worthwhile to recount the sequence of finding the various O-protonated conformers by geometry optimizations at the MP2/6-31G* level. In the initial search the symmetric structures $4n$ and $5m$ were readily found. After this, searches were made of stationary points along the internal rotation path about the C—C bond starting from $\psi_2 = 180°$ ($4n$ or $5m$) toward $\psi_2 = 0°$ (e.g., H4). Along the $4n$ route, with the two amino hydrogens turned away from the C_1—O_4 bond, only $4t$ at $\psi_2 = 135°$ was located. The $5m$ route with the amino hydrogens facing C_1—O_4 led to $5t$, $6m$, and $6t$ at $\psi_2 = 114°$, 93°, and 66°, respectively. Another stationary point found in the general area, but off the $4m$ and $5m$ paths as far as the orientations of the amino hydrogens are concerned is $8t$ at $\psi_2 = 65°$. We note that $6m$ at $\psi_2 = 93°$ is the only energy minimum within the range of $\psi_2 = 180$–$0°$. We also note that any further rotation beyond $\psi_2 = 14°$ toward $\psi_2 = 0°$ would lead to a migration of the proton from the oxygen O_3 to the nitrogen atom. The last to be found was the minimum $4m$ as a distorted structure of $4n$ with a nearly indistinguishable energy difference.

Classification for the O-protonated conformers is not as clear-cut. There are four groups: the lowest energy conformer $4m$ and its transition partners $4n$ and $4t$ connected by C—N rotation (28–29 kcal mol^{-1}); the next two stable minima $5m$ and $6m$ and their transition partners $5t$ and $6t$ via C—C rotation (29–31 kcal mol^{-1}); the higher energy minima $7m$ and $8m$ due to the less stable cis-cis form of $-C(OH)_2^+$ (32–34 kcal mol^{-1}); and the higher barriers $7t$ and $8t$ pertaining to eclipsed conformations in the C—N and C—C rotations originating from $4m$ or $5m$ and $6m$ (33–35 kcal mol^{-1}), respectively.

The relative stability of $4m$ and $5m$ may be explained by the location of the nitrogen lone pair relative to the positively charged $-C(OH)_2^+$ chain. It appears that when the lone pair points toward the C_1—O_4 bond as in $4m$, a greater conjugative attraction arises than when it points away as in $5m$. The smaller C—C—N angle in $4m$ as compared with $5m$ (112° vs. 118° in Table II) is consistent with this explanation. As internal rotation around the C—N or C—C bond moves the nitrogen lone pair or the nitrogen atom away from the planar heavy-atom framework (e.g., $4m \rightarrow 4t$ or $7t$, $5m \rightarrow 5t$ or $7t$, etc.), the conjugative effect decreases while steric repulsion increases for certain 1,4 and 1,5 repulsions between atoms bonded to the N_6, C_2, C_1, and O_3 atoms. These effects are verified by a lengthening of the C—N bond as the dihedral angle ψ_1 or ψ_2 changes from the initial value in $4m$ or $5m$. Finally, the cis-cis form of $-C(OH)_2^+$ is less stable than its cis-trans counterpart ($7m$ vs. $4n$, etc.) because of a very strong 1,5 H···H repulsion.

INTERCONVERSIONS

To explicitly determine the relationship between conformers that lie in close proximity to one another in the PES, IRC calculations[39,40] were performed on all saddle points at the MP2/6-31G* levels. Computation begins with the transition state T and steps downhill on the PES in the forward

direction. The resulting IRC links T to an energy minimum A via the minimum energy path, $T \rightleftarrows A$. The process is repeated for the reverse direction that yields energy minimum B, $T \rightleftarrows B$.

The IRC calculations reveal several pathways for coupled internal rotations about the C—N and C—C bonds but usually with one of the two being more dominant. Table VII summarizes the conversions of the six lower energy minima $(1m, \ldots, 6m)$ to their respective saddle points. (Higher energy barriers for internal rotations around the C—C bond for the N-protonated and C—O bond for O-protonated conformers were not considered.) There are 14 paths listed for the forward conversion $A \rightarrow T$, which are accompanied by changes in the C—N and C—C torsional angles (ψ_1 and ψ_2), and changes in energy with and without correcting for zero-point energies ($\Delta E_e^{\ddagger}$ and $\Delta E_0^{\ddagger}$). The listed values are derived from the MP2/6-311 + G** ΔE_e and ΔE_{ZP} in Table III, assuming that the IRCs identified at the MP2/6-31G* level remain valid at the higher levels.

In the interest of identifying a more complete minimum-energy path (i.e., $A \rightleftarrows T \rightleftarrows B$ where i represents the energy minima at the other end of the barrier) we find the 14 conversions consolidated into four groups of interconversions as shown below.

Group 1: $im \rightleftarrows in \rightleftarrows im'$ for $i = 1, 2, 3$, and 4 and $jm \rightleftarrows jt \rightleftarrows jm'$ for $j = 1, 2$, and 3, where im' and jm' are symmetry equivalent structures of im and ji, respectively. These are paths characterized by rotation about the C—N bond. Based on the small values of $\Delta E_0^{\ddagger}$ at ± 0.2 kcal mol^{-1} for paths 1, 2, 3, 5, and 7, we conclude that the structures in each of the groups $(1m, 1n, 1t)$, $(2m, 2n)$, $(3m, 3n)$, and $(4m, 4n)$ are physically indistinguishable at 0 K due to zero-point vibrations. In fact, the symmetric $1n$–$4n$ become lower in energy than their respective asymmetric partners at 0 K. (A similar conclusion is reached on the basis of the MP2/6-31 + G** values in Table III.) The symmetric in or jt is more likely to be ''observed'' because it is the ''average'' of the symmetry equivalent pair im and im' or jm and jm'. A recent discussion by Godfrey and Brown[24a] on the spectroscopic indistinguishability of a neutral structural pair (the asymmetric IIn vs. the symmetric IIp) applies to the protonated structural pairs encountered here. In view of these findings, the absence of asymmetric structures $1m$–$4m$ from HF/6-31G* optimizations (Table V) no longer bears real physical consequences. The same can be said of the absence of $2m$ and $3m$ from MP2/6-31G* and MP2/6-31G** optimizations (Table III).

The interconversions $2m \rightleftarrows 2t \rightleftarrows 2m'$ and $3m \rightleftarrows 3t \rightleftarrows 3m'$ from paths 4 and 6 are characterized by symmetric double potential wells with very small barriers of 1.1 and 0.6 kcal mol^{-1}, respectively. The latter barrier is comparable to RT at ordinary temperatures and nearly free rotation is expected.

TABLE VI.
Internal Conversions of Protonated Glycine Conformers.

Path	Conversion	ψ_1 (C—N)	ψ_2 (C—C)	$\Delta E_e^{\ddagger}$	$\Delta E_0^{\ddagger}$
(1)	$1m \rightarrow 1n$	$-30 \rightarrow 0$	$11 \rightarrow 0$	0.13	−0.11
(2)	$1m \rightarrow 1t$	$-30 \rightarrow -58$	$11 \rightarrow 0$	0.17	0.18
(3)	$2m \rightarrow 2n$	$-53 \rightarrow -59$	$-166 \rightarrow 180$	0.05	−0.09
(4)	$2m \rightarrow 2t$	$-53 \rightarrow -119$	$-166 \rightarrow 180$	1.52	1.09
(5)	$3m \rightarrow 3n$	$-12 \rightarrow 0$	$-7 \rightarrow 0$	0.01	−0.12
(6)	$3m \rightarrow 3t$	$-12 \rightarrow -58$	$-7 \rightarrow 0$	0.45	0.64
(7)	$4m \rightarrow 4n$	$103 \rightarrow 91$	$-175 \rightarrow 180$	0.02	−0.16
(8)	$4m \rightarrow 4t$	$103 \rightarrow 117$	$-175 \rightarrow 131$	0.44	0.77
(9)	$4m \rightarrow 7t$	$103 \rightarrow 5$	$-175 \rightarrow 163$	4.43	4.25
(10)	$5m \rightarrow 5t$	$-65 \rightarrow -63$	$180 \rightarrow 109$	1.37	1.62
(11)	$5m \rightarrow 7t$	$-65 \rightarrow 5$	$180 \rightarrow 163$	3.82	3.52
(12)	$6m \rightarrow 5t$	$-66 \rightarrow -63$	$95 \rightarrow 109$	0.03	0.01
(13)	$6m \rightarrow 6t$	$-66 \rightarrow -76$	$95 \rightarrow 69$	0.34	0.15
(14)	$6m \rightarrow 8t$	$-66 \rightarrow -1$	$95 \rightarrow 70$	4.41	4.16

Listed for each $A \rightarrow T$ are changes in ψ_1 and ψ_2 (°) and energies (kcal mol^{-1}) with $\Delta E_e^{\ddagger} = \Delta E_e(T) - \Delta E_e(A)$ and $\Delta E_0^{\ddagger} = \Delta E_0(T) - \Delta E_0(A)$ where $\Delta E_0 = \Delta E_e + \Delta E_{ZP}$. The MP2/6-311+G** values for ψ_1, ψ_2, ΔE_e, and ΔE_{ZP} in Tables I–III are used.

Group 2: $4m \rightleftarrows 7t \rightleftarrows 5m$ of paths 9 and 11 characterized by rotation about the C—N bond. The forward and reverse barriers are 4.3 and 3.5 kcal mol^{-1}, respectively.

Group 3: $5m \rightleftarrows 5t \rightleftarrows 6m$ of paths 10 and 12 characterized by rotation about the C—C bond. The forward and reverse barriers are 1.6 and 0.0 kcal mol^{-1}, respectively. In view of the virtually identical $\Delta E_0^{\ddagger}$ for paths 12 and 13, ± 0.2 kcal mol^{-1}, the structures $6m$, $5t$, and $6t$ are taken to be physically indistinguishable.

Group 4: $4m \rightleftarrows 4t$ and $6m \rightleftarrows 8t$ of paths 8 and 14 characterized by rotation about the C—C bond. The forward barriers are 0.8 and 4.2 kcal mol^{-1}, respectively. The end product for path 14 has been positively identified as $1m$, which begins at $8t$ undergoing principally C—C and some C—N rotations to reach a conformation similar to H4 (Fig. 4) before transferring the proton from O to N.[41]

PROTONATION PATHWAYS

By comparing the neutral structures to the protonated structures and taking into consideration the allowed interconversions between protonated conformers, simple protonation pathways are postulated in Table VII. Here the protonation mechanism is represented by the elementary steps that allow the least steric interference to the approach of a proton. The pathways are determined by inspection, which is the adding of the proton to a reactant neutral conformer at the amino N or the carbonyl O site (cf. Fig. 4) and matching the resulting protonated structure to one of the protonated conformers (Figs. 2, 3). (Protonations at the α carbon and hydroxyl oxygen sites do not result in stable molecular ions as shown in previous theoretical studies.[27,28]) With few exceptions, a geometrical match is made that can be a minimum (*im*) or a TS (*it*). Because the end product should be a minimum, an intermediate TS must undergo a conversion to a minimum (Table VI). In the case of the first four energy minima, the symmetric $1n$–$4n$ may be more representative than the asymmetric $1m$–$4m$ because of lower energies at 0 K (Table VI).

TABLE VII.
Protonation Pathways and Relative Gas-Phase Basicities (ΔGB) for Low-Energy Conformers of Glycine.

Pathway	ΔGB	Pathway	ΔGB
	N-Protonations		
I$p \rightarrow 1t \rightarrow 1m$	0.0	V$n \rightarrow 2m$	−0.7
III$n \rightarrow 2m$	−2.3	VI$p \rightarrow 3t \rightarrow 3m$	−2.3
IV$n \rightarrow 1m$	1.4	VIII$n \rightarrow 3m$	−0.7
	O-Protonations		
I$p \rightarrow 8m$	−32.2	IV$n \rightarrow 7m$	−28.4
II$n \rightarrow$ (H4) $\rightarrow 1m$	1.3	V$n \rightarrow 4t \rightarrow 4m$	−24.0
II$n \rightarrow$ (H5) $\rightarrow 3m$	−6.6	V$n \rightarrow 7m$	−27.4
III$n \rightarrow 5m$	−26.9	VI$p \rightarrow 5m$	−22.0
III$n \rightarrow 8m$	−31.5	VIII$n \rightarrow 4m$	−19.2

Minima $1m$–$4m$ may be replaced by satellites $1n$–$4n$ (cf. Table VI). ΔGB (kcal mol^{-1}) is relative to GB of I$p \rightarrow 1m$ that is 202.0 kcal mol^{-1} at the MP2/6-311+G** optimized level after correcting for basis set superposition error.[11] ΔGB is estimated as $-[\Delta G(\mathrm{GlyH}^+) - \Delta G(\mathrm{Gly})]$ based on the values of Table IV.

The few exceptional cases include protonations of IIn and VIIp. VIIp can potentially undergo a sterically favorable O protonation to form the $-C(OH)_2^+$ fragment in a cis-trans or a trans-trans form; but no stationary point in the PES appears to be imminently accessible upon the expected proton entry. There are also the O protonations for IIn where a nonexistent H4 or H5 is postulated as a transient intermediate upon proton entry. But we expect H4 and H5 to end up as the respective N-protonated species $1m$ and $3m$ (*vide supra*).

Based on the relative abundance of the neutral conformers at room temperature, the postulated protonation mechanism, and the relative stability of the protonated conformers, it is unequivocally clear that the most likely protonation product in the gas phase is the global minimum $1m$ (mainly from Ip and some from IIn). There is probably a very small amount of $2m$ (from IIIn). As for the O-protonated species, the Boltzmann distribution law would dictate a totally negligible population. Yet, if a suitable condition prevails for an O protonation, $5m$ (from IIIn) is a more likely product than the lower energy conformer $4m$ (from Vn) because the parent IIIn is more populated than Vn.

In an aqueous solution where glycine exists as a zwitterion and where intermolecular hydrogen bonding plays a significant role in stabilizing different forms of protonated glycine, formations of the N-protonated conformers $1m$, $2m$, and $3m$ and their related structures by C—N internal rotations may become equally probable (cf. Fig. 2). In a peptide the termini consist of the amino and carboxylic groups of glycine. Protonation of a peptide at either terminus or at the amide N or carbonyl O of the peptide bond may introduce internal H-bonded conformations analogous to those depicted

in Figures 2 and 3. Some examples can be found among the protonated diglycine and triglycine conformers calculated at the HF/6-31G* level.[4, 6, 7] Again, in an environment where ionic species are prevalent for intermolecular hydrogen bonding, even the O-protonated conformations may prevail in a peptide.

GAS-PHASE BASICITY

The GB measured by mass spectrometry is the most direct experimental quantity to verify the theoretical model for its computation. On the other hand, theoretical calculations provide the critical structural data for interpreting experimental results and for the design of new experiments. Our recent calculations on the GB and PA of glycine[11] using the highest practical levels yielded 203.5 and 211.1 kcal mol^{-1}, respectively, which is in excellent agreement with the recently updated experimental values of 203.7 and 211.8 kcal mol^{-1} by Hunter and Lias.[36] The theoretical model was based on the assumption that protonation reaction involves only the lowest energy neutral and protonated conformers Ip and 1m. [See eqs. (6)–(8) for reaction (r1).] Contributions to GB from other neutral and protonated conformers were taken into account by way of Boltzmann distributions. Based on the data of Tables V–VII and the foregoing discussion, we now conclude that the theoretical model employed earlier was valid in view of the fact that Ip → 1m is an overwhelmingly dominant protonation pathway at room temperature. Nonetheless, it may be of interest to experimentalists to know what the possible GB values to be observed for the glycine system are. To provide this information, we present relative GB values for all the proposed protonation pathways in Table VII, where ΔGB represents the GB deviation of pathway rk from the most dominant pathway r1. [See eqs. (4) and (5) for reaction (rk).]

We note from the ΔGB of Table VII that the N protonations on conformers higher in energy than Ip yield GB deviations of −2.3 to 1.4 kcal mol^{-1} from GB(r1). These deviations are comparable to the experimental error of ±2.8 kcal mol^{-1} for the glycine GB by mass spectrometry.[42] The O protonations on conformer IIn result in proton transfer to form N-protonated isomers 1m and 3m with small deviations (1.3 and −6.6 kcal mol^{-1}), but all other O protonations bring in GB values some 20–30 kcal mol^{-1} below. Mass spectral measurements on the lower GB values from O protonations of glycine are presently impractical.[42]

Conclusions

We carried out a comprehensive *ab initio* study of the protonated glycine conformers in the gas phase. The structures of the eight energy minima and 12 TSs optimized at the MP2/6-31 + G** and MP2/6-311 + G** levels are expected to be reasonably accurate. Although optimizations at the HF/6-31G*, MP2/6-31G*, and MP2/6-31G** levels led to discrepancies in the conformations of some principal minima, the relative energies were of the right order. The most significant impact made by MP2 optimizations was the elimination of the two O-protonated structures H4 and H5 (Fig. 4) found by HF/6-31G* optimizations. This implied that the attraction from the nitrogen atom for the proton in the arrangement $N\cdots H^+—O$ was so greatly enhanced by electron correlation (via an MP2 optimization) that the proton left the oxygen atom to form the ionic hydrogen bonding $N—H^+\cdots O$.

Relative energies of the nine N-protonated and 11 O-protonated structures were calculated to be 0–10 and 30–40 kcal mol^{-1}, respectively. The relative stability of the protonated species can be satisfactorily explained by the interplay between nonbonded attractive and repulsive interactions. The attractions may be ascribed to hydrogen bonding between the $—NH_3^+$ and —COOH groups in the N-protonated species and to a conjugative effect between the nitrogen lone pair and the $—C(OH)_2^+$ group in the O-protonated species. The repulsions were mainly steric in nature.

The calculated protonated species represent electronic energy minima and saddle points in the low-energy region of the conformational potential energy surface at the MP2/6-311 + G** level of theory. The role of each saddle point as a transition state in a minimum-energy path was identified by separate intrinsic reaction coordinate calculations. The surface was found to be very flat in the immediate vicinities of the asymmetric minima 1m–4m that contain the satellite structure 1n–4n and some TSs. At 0 K the symmetric 1n–4n become lowest in energy after zero-point energy is added. Most paths are internal rotations about the C—N and C—C bonds, but one path corresponds to internal proton transfer. Simple kinetic paths were postulated for the protonation of different glycine conformers. The most stable N-protonated conformer was shown to be the most probable protonation product.

Our findings on the protonated glycine conformers and their interconversions provide important physical data for the study of gas-phase ion chemistry of amino acids and peptides. The protonation pathways suggested for the gaseous glycine may prove useful in the study of biochemical reactions in aqueous solution where protonations and proton transfers are fundamental processes.

Acknowledgments

The computational work support and financial support are gratefully acknowledged. We thank Dr. S. Mark Cybulski for the use of the IBM RS/6000 computer in his laboratory.

References

1. V. N. Balaji and K. Ramnarayan, In *Biological Active Peptides*: *Design*, *Synthesis and Utilization*, W. V. Williams and D. B. Weiner, Eds., Technomic Publishing, Basel, Switzerland, 1991.
2. A. Fersht, *Enzyme Structure and Mechanism*, W. H. Freeman, New York, 1985.
3. K. Biemann and S. A. Martin, *Mass Spectrom. Rev.*, **6**, 1 (1987).
4. K. Zhang, D. M. Zimmerman, A. Chung-Phillips, and C. J. Cassady, *J. Am. Chem. Soc.*, **115**, 10812 (1993).
5. J. W. McKiernan, C. E. A. Beltrame, and C. J. Cassady, *J. Am. Soc. Mass Spectrom.*, **5**, 718 (1994).
6. K. Zhang, C. J. Cassady, and A. Chung-Phillips, *J. Am. Chem. Soc.*, **116**, 11512 (1994).
7. C. J. Cassady, S. R. Carr, K. Zhang, and A. Chung-Phillips, *J. Org. Chem.*, **60**, 1704 (1995).
8. K. A. Jebber, K. Zhang, C. J. Cassady, and A. Chung-Phillips, *J. Am. Chem. Soc.*, **118**, 10515 (1996).
9. S. R. Carr and C. J. Cassady, *J. Am. Soc. Mass Spectrom.*, **7**, 1203 (1996).
10. N. P. Ewing, X. Zhang, and C. J. Cassady, *J. Mass Spectrom.*, **31**, 1345 (1996).
11. K. Zhang and A. Chung-Phillips, *J. Phys. Chem.*, **102**, 3625 (1998).
12. M. Ramek, V. K. W. Cheng, R. F. Frey, S. Q. Newton, and L. Schäfer, *J. Mol. Struct.*, **186**, 12 (1991) and references therein.
13. J. Jensen and M. S. Gordon, *J. Am. Chem. Soc.*, **113**, 7917 (1991).
14. A. G. Császár, *J. Am. Chem. Soc.*, **114**, 9568 (1992).
15. R. F. Frey, J. Conffin, S. Q. Newton, M. Ramek, V. K. W. Chang, F. A. Momany, and L. Schäfer, *J. Am. Chem. Soc.*, **114**, 5369 (1992).
16. C. Chipot, B. Maigret, and L.-L. Rivail, *J. Phys. Chem.*, **96**, 10276 (1992).
17. D. Yu, D. A. Armstrong, and A. Rauk, *Can. J. Chem.*, **70**, 1762 (1992).
18. A. Vijay and D. N. Sathyanarayana, *J. Phys. Chem.*, **96**, 10735 (1992).
19. A. A. Bliznyuk, H. F. Schaefer III, and I. J. Amster, *J. Am. Chem. Soc.*, **115**, 5149 (1993).
20. C.-H. Hu, M. Shen, and H. F. Schaefer III, *J. Am. Chem. Soc.*, **115**, 2923 (1993).
21. V. Barone, C. Adamo, and F. Lelj, *J. Chem. Phys.*, **102**, 364 (1995).
22. I. D. Reva, A. M. Plokhotnichenko, S. G. Stepanian, A. Y. Ivanov, E. D. Radchenko, G. G. Sheina, and Y. P. Blagoi, *Chem. Phys. Lett.*, **232**, 141 (1995), [Erratum] **235**, 617 (1995).
23. A. G. Császár, *J. Mol. Struct.*, **346**, 141 (1995).
24. (a) P. D. Godfrey and R. D. Brown, *J. Am. Chem. Soc.*, **117**, 2019 (1995); (b) P. D. Godfrey, R. D. Brown, and F. M. Rodgers, *J. Mol. Struct.*, **376**, 65 (1996).
25. J. J. Neville, Y. Zheng, and C. E. Brion, *J. Am. Chem. Soc.*, **118**, 10533 (1996).
26. M. Ramek, F. A. Momany, D. M. Miller, and L. Schäfer, *J. Mol. Struct.*, **375**, 189 (1996).
27. L. R. Wright, R. F. Borkman, and A. M. Gabrielll, *J. Phys. Chem.*, **86**, 3951 (1982).
28. S. Bouchonnet and Y. Hoppilliard, *Org. Mass Spectrom.*, **27**, 71 (1992).
29. F. Jensen, *J. Am. Chem. Soc.*, **114**, 9533 (1992).
30. D. Yu, A. Rauk, and D. A. Armstrong, *J. Am. Chem. Soc.*, **117**, 1789 (1995).
31. See a description of the *ab initio* methods and notations in W. J. Hehre, L. Radom, P. v. R. Schleyer, and J. A. Pople, *Ab Initio Molecular Orbital Theory*, Wiley, New York, 1986.
32. See, for example, M. J. Frisch, G. W. Trucks, H. B. Schlegel, P. M. W. Gill, B. G. Johnson, M. A. Robb, J. R. Cheeseman, T. Keith, G. A. Petersson, J. A. Montgomery, K. Raghavachari, M. A. Al-Laham, V. G. Zakrzewski, J. V. Ortiz, J. B. Foresman, J. Cioslowski, B. B. Stefanov, A. Nanayakkara, M. Challacombe, C. Y. Peng, P. Y. Ayala, W. Chen, M. W. Wong, J. L. Andres, E. S. Replogle, R. Gomperts, R. L. Martin, J. D. Fox, J. S. Binkley, D. J. Defrees, J. Baker, J. J. P. Stewart, M. Head−Gordon, C. Gonzalez, and J. A. Pople, *Gaussian 94*, Gaussian, Pittsburgh, PA, 1995.
33. J. A. Pople, A. P. Scott, M. W. Wong, and L. Radom, *Isr. J. Chem.*, **33**, 345 (1993).
34. A. P. Scott and L. Radom, *J. Phys. Chem.*, **100**, 16502 (1996).
35. See references cited in (a) D. E. Wood and T. H. Dunning, Jr., *J. Chem. Phys.*, **103**, 4572 (1995); (b) A. K. Wilson and T. H. Dunning, Jr., *J. Chem. Phys.*, **106**, 8718 (1997).
36. E. P. Hunter and S. G. Lias, *J. Phys. Chem. Ref. Data*, to appear; see also W. G. Mallard and P. J. Linstrom, Eds., *NIST Standard Reference Database Number 69*, National Institute of Standards and Technology, Gaithersburg, MD, 1997.
37. A. G. Császár, *J. Phys. Chem.*, **100**, 3541 (1996).
38. (a) G. A. Jeffrey and W. Saenger, *Hydrogen Bonding in Biological Structures*, Springer−Verlag, New York, 1991; (b) T. Steiner and W. Saenger, *Acta Crystollogr.*, **B48**, 819 (1992).
39. K. Fukui, *Acc. Chem. Res.*, **14**, 363 (1981).
40. C. Gonzalez and H. B. Schlegel, *J. Chem. Phys.*, **90**, 2154 (1989), **94**, 5523 (1990).
41. K. Zhang and A. Chung-Phillips, unpublished results.
42. C. J. Cassady, private communication, 1997.